WHAT IS CLAIMED IS:

1. A method of preparing a species-specific nucleic acid database comprising:

- selecting from a species-non-specific nucleic acid database species-specific nucleic acids comprising coding sequences;

- selecting from a species-non-specific nucleic acid database species-specific nucleic acids comprising noncoding sequences;

- selecting from the coding sequences those sequences that are 3′-compete or 3′-coding biased, wherein 3′-coding biased sequences comprise 5′-partial sequences having desirable characteristics;

- selecting from the noncoding sequences those sequences that include poly-A tails or are derived from sequences that include poly-A tails;

- reducing redundancy in selected sequences;

- comparing sequences comprising unannotated sequences to a collection of sequences comprising annotated coding sequences and selecting those sequences satisfying a threshold of similarity; and

- collecting all selected sequences.

2. The method according to claim 1, wherein the species-specific nucleic acid database is an equine-specific nucleic acid database.

3. The method according to claim 1, wherein the species-non-specific nucleic acid database is GenBank.

4. An array comprising a plurality of oligonucleotide probes designed to be complementary to and hybridize under stringent conditions with a gene listed in one of Tables 33, 35, or 37.

5. The array according to claim 4, wherein the array consists of less than 100 probes that are complementary to genes not listed in Tables 33, 35, or 37.

6. The array according to claim 4, wherein the array is designed for diagnosis of disease.

7. The array according to claim 6, wherein the array is designed for diagnosis of equine or canine disease.

8. The array according to claim 4, wherein the array comprises at least one gene or sequence shown in Table 9 or 10, and wherein in the array is designed for diagnosis of disease in any tissue of any animal.

9. An array comprising a plurality of oligonucleotides, wherein:

a) the oligonucleotides are chosen from the nucleic acid sequences shown in Tables 34, 36, or 38, and wherein the array comprises 10 or more of said oligonucleotides; or

b) the oligonucleotides comprise nucleotide probes designed to be complementary to, or hybridize under stringent conditions with, 10 or more nucleic acid sequences shown in Tables 34, 36, or 38.

10. The array according to claim 9, wherein the oligonucleotides comprise nucleotide probes designed to be complementary to, or hybridize under stringent conditions with, 1000 or more nucleic acid sequences shown in Table 6.

11. The array according to claim 10, wherein the oligonucleotides comprise nucleotide probes designed to be complementary to, or hybridize under stringent conditions with, 2000 or more nucleic acid sequences shown in Table 6.

12. The array according to claim 11, wherein the oligonucleotides comprise nucleotide probes designed to be complementary to, or hybridize under stringent conditions with, 3000 or more nucleic acid sequences shown in Table 6.

13. A method for populating a database of species-specific nucleic acid sequences, comprising:

querying a database of nucleic acid sequences to identify nucleic acid sequences associated with a subject species;

processing the identified sequences to create a first subset containing coding sequences and a second subset containing non-coding sequences;

dividing the first subset into a plurality of DNA sequences, if present, and a plurality of mRNA sequences;

processing the plurality of DNA sequences to derive a plurality of virtual mRNA sequences;

dividing the plurality of mRNA sequences into a plurality of complete and mRNA 3' partial sequences, and a plurality of mRNA 5' partial sequences;

processing the plurality of mRNA 5' partial sequences to identify a subset of mRNA 5' partial sequences, each member of the subset satisfying a threshold level of completeness;

identifying members of the second subset containing non-coding sequences that correlate with at least one known coding sequence of at least one species other than the subject species; and

combining the plurality of virtual mRNA sequences, the plurality of complete and MRNA 3' partial sequences, the subset of mRNA 5' partial

sequences, and the identified correlated sequences to create the database of species-specific nucleic acid sequences.

14. The method according to claim 13, wherein the step of identifying includes comparing each member of the second subset to each member of a database containing annotated human nucleic acid sequences.

15. The method according to claim 13, wherein the step of identifying includes comparing each member of the second subset to each member of a database containing annotated human and mouse nucleic acid sequences.

16. The method according to claim 15, wherein the database containing annotated human and mouse nucleic acid sequences is derived from the database of nucleic acid sequences.

17. The method according to claim 13, further comprising eliminating duplicates within the database of species-specific nucleic acid sequences.

18. The method according to claim 13, further comprising populating the database of species-specific nucleic acid sequences with selected species-specific virus definitions.

19. The method according to claim 13, further comprising verifying that each of the identified correlated sequences is represented in sense format.

20. A method of identifying changes in gene expression with time, comprising assaying a biological sample with the microarray according to claim 4, repeating the assay after a period of time has elapsed, and comparing the results.

21. A method of detecting or monitoring a disease chosen from osteoarthritis, joint inflammation, neurological diseases, developmental

orthopedic diseases, laminitis, and the general condition of stress, comprising testing a biological sample on a microarray according to claim 4 for the presence of a genetic marker associated with the disease being tested for.

22. The method according to claim 21, wherein the neurological disease is equine protozoal myelitis.

23. A method of detecting or monitoring an infectious disease chosen from herpesvirus-2 and equine protozoal myelitis caused by sarcocystis neurona or sarcocystis neurospora, comprising testing a biological sample on a microarray according to claim 4 for the presence of a genetic marker associated with the disease being tested for.